

## Education Select Committee

### 2012 GCSE English Response from Ofqual



#### Introduction

This memorandum provides the Select Committee with information responding to the questions submitted to us.

The questions are all detailed individually with the response following where possible. Where we needed to include additional tables and charts, these are available as appendices.

In order to provide a response we collected a considerable amount of information from exam boards. To keep our data request reasonable within the timescales involved, we requested detailed candidate data for the new GCSE English and GCSE English language qualifications only. Candidate level data was not requested for the established GCSE in English literature. If on reviewing the information we have provided, the Select Committee would like us to make an additional request to exam boards, then we will do this.

We would welcome the opportunity to discuss our responses with the Select Committee.

A handwritten signature in black ink, appearing to be 'Glenys Stacey'.

Glenys Stacey  
Chief Regulator  
18th October 2012

## Market share

1. **Please supply a breakdown of market share (in percentage and numbers of candidates) for GCSE English, GCSE English language and GCSE English literature across the exam boards for candidates in England in 2012 and for the equivalent syllabuses in 2011.**

### Response

QCA revised the GCSE English and English literature subject criteria in 2008/9 for qualifications to be taught from September 2010 and for first qualification awards in June 2012.

Previously there were two qualifications – English and English literature. Students in maintained schools in England had to enter English and could then choose whether to enter English literature as well. The new syllabuses for 2010 offered a choice of three qualifications – English, English language and English literature. Students in maintained in England schools could choose to enter English (oneGCSE), or to enter English language and English literature (twoGCSEs).

The previous syllabuses had been remarkably stable, in terms of the content (the National Curriculum Programme of Study for English) and the nature of the assessment. There are significant differences between the legacy GCSE English specifications and the new GCSE English specifications.

- The qualification has changed from one subject studied by all to two subjects available for study: GCSE English and the new GCSE English language.
- The revised English GCSE has a different structure from the legacy qualification and now has common units available across the two qualifications.
- The subject criteria are now more detailed and extensive.
- The qualification has changed from linear assessment – all assessments had to be taken at the end of the course at 60 per cent (two exams), to unitised assessment – candidates could enter assessments throughout the course and re-sit those entered early. Specifications must have allocated a weighting of 40 per cent to external assessment. A terminal rule of 40 per cent was introduced.
- The qualification has changed from teacher-marked assessment (coursework) worth 40 per cent of the qualification, to teacher-marked assessment (controlled assessment) worth 60 per cent of the qualification.
- There are differences in weightings and paper durations.

- There is a new requirement to “participate in a range of real-life contexts in and beyond the classroom”.
- There now must be a weighting of 45 55 per cent for the “functional elements of English”. The extent of the weighting indicates the importance of this change.

We will explain these in more detail in our final report.

The tables below show the breakdown in market share in England and for all UK by awarding organisation. AQA have over half of the market share for GCSE English, GCSE English language and GCSE English literature.

Given that the English language and English specifications are new for 2012, there are no strictly equivalent specifications in 2011. In 2011, AQA English A had over 60 per cent of the candidature. Equivalent specifications were available in English Literature. In 2011 AQA English literature A had the largest market share with over 68 per cent of candidates in England. CCEA had no candidates in England in 2012.

## 2012 GCSE English

AO	Spec title	Spec code	England		All UK	
			Candidates	Percentage	Candidates	Percentage
AQA	English	4702	94486	56.9%	98407	57.7%
CCEA	English	G9310	0	0.0%	198	0.1%
Edexcel	English	2EH01	22242	13.4%	22428	13.2%
OCR	English	J350	10868	6.5%	10925	6.4%
WJEC	English	4190SA	38419	23.1%	38553	22.6%
			166015	100.0%	170511	100.0%

## 2012 GCSE English language

AO	Spec title	Spec code	England		All UK	
			Candidates	Percentage	Candidates	Percentage
AQA	English Language	4707	286536	63.6%	292402	57.7%
CCEA	English Language	G9290	0	0.0%	14823	2.9%
Edexcel	English Language	2EN01	45398	10.1%	46209	9.1%
OCR	English Language	J355	34266	7.6%	35569	7.0%
WJEC	English Language	4170SA	84261	18.7%	117909	23.3%
			450461	100.0%	506912	100.0%

## 2011 GCSE English

			England		All UK	
AO	Spec title	Spec code	Candidates	Percentage	Candidates	Percentage
AQA	English A	3702	369875	63.7%	375836	58.4%
	English B (mature)	3703	15712	2.7%	15858	2.5%
	English B	3701	26255	4.5%	27284	4.2%
CCEA	English	G29	120	0.0%	19743	3.1%
Edexcel	English	2731	1685	0.3%	1685	0.3%
	English A	1203	14621	2.5%	14922	2.3%
	English B	1204	5799	1.0%	5979	0.9%
OCR	English	1900	32901	5.7%	34269	5.3%
WJEC	English	15001,	113803	19.6%	148338	23.0%
		15002,				
		15101,				
		15102				
			580771	100.0%	643914	100.0%

Data from Joint Council for Qualifications (JCQ) dataset, excludes “ABS” results

## 2012 English literature

AO	Spec title	Spec code	England		All UK	
			Candidates	Percentage	Candidates	Percentage
AQA	English Literature	4712	269373	62.9%	275379	58.7%
CCEA	English Literature	G9300	0	0.0%	7390	1.6%
Edexcel	English Literature	2ET01	42528	9.9%	43130	9.2%
OCR	English Literature	J360	34224	8.0%	35214	7.5%
WJEC	English Literature	4200SA	82170	19.2%	108004	23.0%
			428295	100.0%	469117	100.0%

## 2011 English literature

AO	Spec title	Spec code	England		All UK	
			Candidates	Percentage	Candidates	Percentage
AQA	English Literature B	3711	23229	5.2%	25066	5.1%
	English Literature A	3712	306237	68.2%	309672	63.3%
CCEA	English Literature	G30	118	0.0%	8157	1.7%
Edexcel	English Literature	1213	15101	3.4%	15515	3.2%
OCR	English Literature	1901	22998	5.1%	23874	4.9%
WJEC	English Literature A	15301, 15302	80236	17.9%	105629	21.6%
	English Literature B	15303, 15304	1108	0.2%	1393	0.3%
			449027	100.0%	489306	100.0%

Data from JCQ dataset, excludes absent results

## Variations in performance in summer 2012

2. **What have you learnt since your initial report about the pattern of results across schools and colleges? What more do you know about whether some types of school/college been disproportionately affected and the reasons for this? What more do you know about the types of schools/colleges that have done better than expected or that have seen improvements in their English results and the reasons why?**
  - a) For each exam board, how many centres had a 5, 10 or 15+ percentage point change in outcomes compared with summer 2011, and how many of these centres had improved outcomes and how many had lower outcomes?
  - b) Please send analyses against as many variables as you have to try to establish patterns and the reasons for these patterns, both for centres which did very well and vice versa: e.g. type, region, size, upper age, percentage foundation tier/percentage higher tier, most recent Ofsted rating, GCSE Maths results, standard deviation of previous years' English results, proportion of re-sit candidates, proportion of January/June entrants.
  - c) Please supply a sample of statistical outliers (in both directions) and establish in depth with the centre what happened and why.

## Response

See attached documents:

Question 2a – Pattern of change across awarding organisations

Question 2b – Understanding reasons for patterns

- a) The table attached shows the pattern of change across awarding organisations between 2011 and 2012 and 2010 and 2011 (data was limited to centres with more than 50 candidates). To summarise, 67.5 per cent (1,747) of centres saw a change of five or more percentage points in their English A\* to C results from 2011 (whether upwards or downwards). Of these, 30.1 per cent (780) of centres saw an improvement in the proportion of students achieving A\* C in English and 37.4 per cent (967) experienced a decrease. This did vary across awarding organisations.

By comparison, 51.9% of centres saw a change of 5 or more percentage points within A\*-C results between 2010 and 2011. Of these 26.8% saw an improvement and 25.2% experienced a decrease.

- b) Analyses have been provided for a number of variables – see attached tables. These include:
- Ofsted rating
  - proportion of re-sit candidates
  - whether the centre was single or mixed gender
  - region of centre
  - whether centre has a sixth form
  - type of centre
  - upper age of centre
  - number of GCSE English candidates
  - entry patterns (January/June entrants)

The analyses showed that there were relationships with all of the variables analysed and the change in centre results at the percentage A\* C boundary. To understand whether the patterns seen were unique to 2012 compared to 2011, analyses were also conducted for 2011 results compared to 2010 where appropriate.

The results showed the following:-

- The North East and North West regions had the most variability in 2012 versus 2011. They also had the most variability in 2011 versus 2010 but to a lesser extent.
- Seventy-three per cent of colleges saw a decline in results of more than 15 per cent in their A\* C. In 2011 versus 2010 they were the centre type with the most variability but only 14.5 per cent of their centres had results with more than a 15 per cent decline.
- Independent schools had little variability in their results in 2012 versus 2011 with 73 per cent of their centres having +/- 5 per cent in their percentage of A\* C results when compared with 2011. However, this is still lower than in 2011, where 91 per cent of centres were within +/- 5 per cent of their percentage of A\* C results in 2010.
- Centres with a statutory highest age of 19+ years were the most affected by a decline in percentage of A\* C results in 2012 versus 2011.
- Centres with 50-99 candidates taking GCSE English (the smallest cohort) were the centres that were most affected, with a decline in their percentage of A\* C results between 2012 and 2011. This was also so in 2011 versus 2010 but to a lesser extent.
- Centres with an outstanding Ofsted rating were less likely to see year-on-year increases in the percentage of A\* C results.
- Centres that saw a year-on-year increase in results had a higher proportion of re-sit candidates.
- Centres that saw a year-on-year increase in results had a higher proportion of candidates taking at least one unit before June 2012.

#### c) Qualitative feedback from centres

As part of our investigation into GCSE English we commissioned a research company to carry out interviews with a range of schools and colleges across England.

We aim to complete around 100 interviews before publication of our findings later this month. At 17th October 2012, 98 centre interviews had been conducted. Seven of those interviews conducted to date were with centres that we have defined as statistical outliers (outside 1.96 standard deviations), as shown in the table below.

Centre	% change in GCSE English A* C pass rate between 2011 and 2012
A state maintained girls' school in Manchester	- 34.48
A state maintained boys' school in Liverpool	- 62.82
A co-educational voluntary-aided comprehensive school in Cheshire	- 39.74
A co-educational comprehensive school in Milton Keynes	- 39.55
A co-educational academy in Essex	+ 34.71
A co-educational secondary school in Berkshire	+ 32.87
A sixth form college in West London	- 42.88

Two of these centres experienced a large increase in their GCSE English A\* C pass rate this year when compared with last year, and **five** of the centres had a much lower pass rate this year. The exam board used by all **seven** centres was AQA.

The interviews were a combination of face-to-face and telephone interviews. A range of staff at each centre were interviewed, depending upon their availability on the day. Interviewees included headteachers, deputy heads, heads of English, exams officers and teachers.

In a letter sent to the centres invited to take part in we stated that “the interviews will be carried out in confidence and the views of individual centres will not be identified when we report back our findings.”

#### What entry approaches were used?

The centres adopted a range of different entry timings. Some entered the full cohort for the written examination at the end of Year 10 and then sat the written exam again at the end of Year 11. Others took the exam only at the end of Year 11. The timing of controlled assessments also varied: some centres submitted controlled assessment at the end of Year 10, with the opportunity for re-sits in Year 11. Other centres carried out most of the controlled assessment at the end of Year 11. At the sixth form college all students completed a one-year course to re-sit their English GCSE. Across the **seven** centres there were no entry strategies unique to the centres that experienced

an increase in their pass rates, or strategies unique to the centres with lower pass rates.

The centres also used a range of subject entry approaches: some entered all students for the single award English qualification, some entered all students for the English language and English literature GCSEs, and other centres chose a subject entry strategy based upon the ability of each student. These centres also reflected a range of strategies in terms of entering students for the Higher and Foundation tiers of the written examination.

#### Which students were affected?

The main area of concern across the five centres reporting lower pass rates this year was students who had been predicted a C grade and received a D grade. A minority of centres also reported a decrease in the proportion of students achieving A\* and A grades.

The **two** centres that achieved an improved pass rate in English GCSE in 2012 compared with 2011 were also disappointed with their A\* C pass rates because they were lower than the centres had predicted.

#### What happened?

All the centres stated that they had worked hard to understand the nature of the new specifications for GCSE English, and were confident that they had prepared thoroughly and effectively for the demands of the new specifications. All **seven** centres invested time in attending meetings and seminars organised by the exam boards and studying the supporting documentation provided by the exam board (in this case AQA).

Across the centres, the English teaching teams had a track record of predicting students' final GCSE grades accurately. They felt as confident in their predictions this year as they had felt about their predictions in previous years. A number of centres mentioned that even though the English qualifications were new this year, "years of teaching" had equipped teachers with a "feel" for what a C grade looked like. Where centres mentioned the feedback that they had received on their moderation of controlled assessment, the moderation had been positive about the accuracy of the centres' marking.

Each centre had a different perspective on "what happened" with the GCSE English qualifications this year. The concerns expressed by the centres varied greatly but fell into three main categories:

- The extent to which the grade boundaries changed took centres by surprise. All seven centres used the January grade boundaries as the basis for their



predictions. Some centres said that they would have acted differently had they known the grade boundaries would change.

- The seven centres expressed a number of concerns about the delivery of controlled assessments:
  - The structure of controlled assessment did not support effective teaching and learning.
  - Some centres believed that across centres there was a high degree of variation in interpretation of the rules around controlled assessment, for example around the extent to which students should be supervised and should be allowed to use notes.
  - The exam board's moderation of teachers' marking of controlled assessment did not effectively correct overly severe or generous marking.
  - Some centres felt positive about controlled assessment and believed it had worked well for them.
- There were relatively few criticisms of the exam, but two areas of concern were expressed:
  - At two hours 15 minutes, the written exam was too long for less able students.
  - The Foundation paper included a question about a radio script, which was unexpected and which centres believed students would struggle to answer.

**3. You mentioned in oral evidence that candidates at the C/D borderline and in FE colleges seem to have been particularly affected. What more do you know about groups of candidates that have been most affected and the reasons for this?**

**Response**

There are three different ways at looking at this question. They are:

- a) Are there different characteristics of centres that have seen variation between 2012 and 2011?
- b) Does the variation occur at all the grade boundaries?  
  
See attached document – Candidates affected by the C/D borderline.
- c) Is there any impact by age and centre type?

a) Question 2b addresses this first question. We saw in the analyses that there were different characteristics that affected the variability in results. An example of this is that 73 per cent of colleges have seen a decline in results of more than 15 per cent in their A\* C percentage.

b) The six tables attached show the centre variability between 2012 and 2011 and between 2011 and 2012 at different grade boundaries: A\* B, A\* C and A\* D. The tables illustrate that there was greater variability in the percentage of candidates at centres achieving A\* B and A\* C grades than those achieving A\* D grades. The variability between 2011 and 2010 showed a similar pattern but it was less prominent.

The third table shows the difference in percentage of centres achieving A\* C in English between 2012 and 2011 by awarding organisation. Across all the awarding organisations, 67.4 per cent of centres saw more than a 5 per cent change in the proportion achieving A\* C in English. At the A\* D level this figure fell to 44.7 per cent. This would suggest a movement of a number of candidates from C to D grade.

There is anecdotal information from our consultation with centres that would tend to support this. We asked schools and colleges whether there were any particular students affected.

Where centres were able to identify groups of students affected, almost all suggested that the impact had been at the C/D boundary:

“There is a clear ‘bubble’ around the C/D grade boundary where the marks have been altered most.”

“The bulk of the difference between actual and predicted grades was across Cs and Ds with ‘solid’ Cs getting a D.”

“If the school had entered their C/D candidates into Foundation English Language they would have got more Cs.”

A minority of centres stated that the impact had been greatest on “vulnerable” students and those claiming free school meals; *“It was the most vulnerable children who were affected.”*

c) The three tables below show the centre variability between 2012 and 2011 at the C/D boundary by candidate characteristics: gender, age and centre type.

The table below shows the difference in percentage of candidates achieving A\*-C between 2012 and 2011 in individual centres.

### Pattern of change in centres between 2012 to 2011 in A\*-C results, by gender of candidates

Difference band (Difference in % A*-C 2012 vs 2011)	Number of	Male candidate	% of Male candidate	Female candidate	% of Female
More than 15% decrease	56317	29302	12.9	27015	12.2
Between 10% and 15% decrease	43451	22432	9.9	21019	9.5
Between 5% and 10% decrease	62988	30727	13.6	32261	14.5
Between 5% decrease and 5% increase	141113	69466	30.7	71647	32.3
Between 5% and 10% increase	45639	23240	10.3	22399	10.1
Between 10% and 15% increase	33336	17399	7.7	15937	7.2
More than 15% increase	65617	33770	14.9	31847	14.3
Total	448461	226336		222125	

**Notes:**

Only centres with 50 or more candidates for 2011 and 2012 with an individual awarding organisation included

If a centre has entered more than 50 candidates with multiple awarding organisations for 2011 and 2012, results by each awarding organisation used will be counted

The table above shows there was a similar proportion of males and females at all levels of centre variation, suggesting there was little impact of gender on centres' results.

Below is the differences in percentage of candidates achieving A\*-C by age. Note, this is the age of the candidate at the start of the academic year.

### Pattern of change in centres between 2012 to 2011 in A\*-C results, by age of candidates

Difference band (Difference in % A*-C 2012 vs 2011)	More than 15% decrease	Between 10% and 15% decrease	Between 5% and 10% decrease	Between 5% decrease and 5% increase	Between 5% and 10% increase	Between 10% and 15% increase	More than 15% increase	Total
Aged 12 & 13	147	288	273	70	37	146	13	974
% Aged 12 & 13	15.1	29.6	28.0	7.2	3.8	15.0	1.3	
Aged 14	6325	3265	4082	6268	2680	2312	3370	28302
% Aged 14	22.3	11.5	14.4	22.1	9.5	8.2	11.9	
Aged 15	40661	38176	55671	129943	42026	30116	60324	396917
% Aged 15	10.2	9.6	14.0	32.7	10.6	7.6	15.2	
Aged 16	3986	1038	2028	3784	732	620	1281	13469
% Aged 16	29.6	7.7	15.1	28.1	5.4	4.6	9.5	
Aged 17+	5194	684	934	1047	164	142	629	8794
% Aged 17+	59.1	7.8	10.6	11.9	1.9	1.6	7.2	
Total candidates	56317	43451	62988	141113	45639	33336	65617	448461
% of Total candidates	12.6	9.7	14.0	31.5	10.2	7.4	14.6	

**Notes:**

Age is at start of the academic year

Only centres with 50 or more candidates for 2011 and 2012 with an individual awarding organisation included

If a centre has entered more than 50 candidates with multiple awarding organisations for 2011 and 2012, results by each awarding organisation used will be counted

When comparing the impact of age at all levels of centre variation, centres with a decrease in A\*-C grades, showed a higher proportion of pupils aged 12 and 13 (around 29% were in centres which saw a decrease of 10% and 15%) and in 14 year olds (22% were in centres seeing more than 15% decrease). These pupils are most likely early entrants to the GCSE. In centres where there is more than a 15% decrease at A\*-C, there is a high proportion of pupils aged 17 plus (59%), in comparison to any other bands. These candidates are most likely to be based in further education establishments, which we know were the most affected.

## Pattern of change in centres between 2012 to 2011 in A\*-C results, by type of centre

Difference band (Difference in % A*-C 2012 vs 2011)	More than 15% decrease	Between 10% and 15% decrease	Between 5% and 10% decrease	Between 5% decrease and 5% increase	Between 5% and 10% increase	Between 10% and 15% increase	More than 15% increase	Total
Secondary Comprehensive (CVAC)	27167	24396	37482	66262	28090	20786	36145	240328
% of Secondary Comprehensive (CVAC)	11.3	10.2	15.6	27.6	11.7	8.6	15.0	
Secondary Selective (CVAC)	0	117	689	10070	460	0	0	11336
% of Secondary Selective (CVAC)	0.0	1.0	6.1	88.8	4.1	0.0	0.0	
Secondary Modern (CVAC)	3435	1074	2094	3639	355	886	2835	14318
% of Secondary Modern (CVAC)	24.0	7.5	14.6	25.4	2.5	6.2	19.8	
Secondary Comprehensive (F)	3399	4416	6499	11902	3339	2310	5434	37299
% of Secondary Comprehensive (F)	9.1	11.8	17.4	31.9	9.0	6.2	14.6	
Secondary Selective (F)	0	0	169	3050	0	190	0	3409
% of Secondary Selective (F)	0.0	0.0	5.0	89.5	0.0	5.6	0.0	
Secondary Modern (F)	543	115	1025	842	134	1102	812	4573
% of Secondary Modern (F)	11.9	2.5	22.4	18.4	2.9	24.1	17.8	
Independent	364	397	1891	9471	411	186	512	13232
% of Independent	2.8	3.0	14.3	71.6	3.1	1.4	3.9	
Further Education Est.	4338	594	661	458	0	0	0	6051
% of Further Education Est.	71.7	9.8	10.9	7.6	0.0	0.0	0.0	
Sixth Form College	2320	0	370	513	50	0	482	3735
% of Sixth Form College	62.1	0.0	9.9	13.7	1.3	0.0	12.9	
Tertiary College	1269	212	456	52	0	0	156	2145
% of Tertiary College	59.2	9.9	21.3	2.4	0.0	0.0	7.3	
Other (inc. private candidates)	187	319	133	832	251	0	483	2205
% of Other (inc. private candidates)	8.5	14.5	6.0	37.7	11.4	0.0	21.9	
Non-NCN Classified	78	174	392	532	99	0	272	1547
% of Non-NCN Classified	5.0	11.2	25.3	34.4	6.4	0.0	17.6	
City Academy	13217	11637	11127	33490	12450	7876	18486	108283
% of City Academy	12.2	10.7	10.3	30.9	11.5	7.3	17.1	
Total candidates	56317	43451	62988	141113	45639	33336	65617	448461
% of Total candidates	12.56	9.69	14.05	31.47	10.18	7.43	14.63	

### Notes:

(CVAC) Community, Voluntary Aided and Controlled

(F) = Foundation

(NCN) = Centre does not have a national centre number

Only centres with 50 or more candidates for 2011 and 2012 with an individual awarding organisation included

If a centre has entered more than 50 candidates with multiple awarding organisations for 2011 and 2012, results by each awarding organisation used will be counted

Candidates in centres with the least variability were most likely from independent schools and secondary selective schools. Candidates in further education establishments, sixth form colleges and tertiary colleges were more likely to show more than a 15% decrease in their A\*-C results. Just under a quarter (24%) of candidates in secondary modern centres saw an overall decline in the centre by more than 15% and almost 20% of candidates in the same centre type saw more than a 15% increase in A\*-C grades overall.

## 4. Is there a link between the variation in performance at school/college-level and the comparable outcomes approach?

### Response

The comparable outcomes approach is not used at centre level. It is an approach to setting grade standards for the whole cohort.

We believe the variation at centre level is due to a combination of several factors which play out differently in each centre:

- historically, English results were more stable than for any other subject
- uncertainty caused by the changes in the qualification
- a “route effect” which meant (generally) those taking (and re-sitting) units did better than those who entered all the units at the end
- moderation which was not effective given some of these other factors.

We will explain this in much more detail in our final report.

**5. How does the school-level variation seen in summer 2012 compare with the school-level variation that occurred in 2010 and 2011?**

**Response**

See attached tables – pattern of change at A\* C borderline

The tables attached compare the variation at A\* C level for school centres between 2011 and 2012. The analysis has the following caveats, however. This data has been taken from Edubase (a Department for Education database of all educational establishments across England and Wales) and does not include centres in Northern Ireland. There has been some significant change in the awarding organisations chosen by centres between 2011 and 2012; possibly some centres have chosen to switch because of the introduction of a new qualification. To control for the impact of this, we have only made comparisons between centres which have remained with the same awarding organisations over the two years. This has affected numbers for Edexcel in particular, which saw a significant gain in its centres for GCSE English, but due to the exclusion of switching, has very low numbers of centres in the table below.

The tables indicate that, across all awarding organisations, the school level variation between 2012 and 2011 was greater than that seen between 2011 and 2010. In 2012, 68 per cent of centres saw a variation of more than 5 per cent in their GCSE English A\* C results, compared with 52 per cent in 2011. This variation is both up and down, with 36.3 per cent of centres experiencing a drop in results of more than 5 per cent and 31.9 per cent with an increase of more than 5 per cent. In 2011 the corresponding figures were 25.0 per cent and 27.0 per cent respectively. Greater variation may have been anticipated in 2012 given the introduction of the new specifications.

In addition the final table shows the mean and standard deviations of the difference in A\* C results for 2012 versus 2011 and 2011 versus 2010 for each exam board

and at a total level. The standard deviation is a measure of variability of the year on year results at the % A\* – C grade. The table clearly shows that the standard deviation has increased considerably from 10 per cent to 15 per cent in 2012. This increase is seen across all exam boards.

The results of this analysis should be treated with caution given the caveats we have set out above.

**6. Have concerns been raised with you about results in other GCSE subjects? If so, please send details.**

**Response**

See attached document – ASCL concerns regarding GCSE Mathematics and letter from ASCL to Ofqual

Between 24th August 2012 and 5th October 2012, our help desk received 617 queries relating to marking, results and the awarding of grades for GCSEs. This compares with 156 queries on the same topics for the period August 2011 – July 2012 and 32 in August 2010 – July 2011.

The vast majority (84 per cent) of these queries were related to English in 2012 compared with 14 per cent for 2011/12 and 12.5 per cent for 2010/11. More specifically, 99 per cent of these queries for English in 2012 related to the setting of grade boundaries, compared with less than 5 per cent for the previous two years. If we removed English from the list of queries made on these topics, the volume actually decreased between last year and this year.

Queries relating to marking, results and the awarding of grades for mathematics also rose this year. This is not surprising as it is a new specification. There were 27 overall for August 2012, 6 for August 2011 – July 2012 and none in August 2010 – July 2011. Most of the complaints relating to mathematics were about setting of grade boundaries, 88 per cent in 2012 and 83 per cent in 2011. Additionally, the Association of School and College Leaders (ASCL) raised concerns related to the awarding of mathematics in 2012 that were forwarded to us. That communication is attached.

In response to the queries to our help desk and the ASCL concerns we reviewed awarding data, observer reports generated as part of our qualification scrutiny programme and other documentation and have not identified that any mathematics awards were unsafe. We have also asked the exam boards to respond to the matters that ASCL raise and we are awaiting their comments before responding formally to ASCL. We believe that the concerns raised reflect normal variation for the first awards of a new specification and that the attention focused on English made some schools/colleges more likely to raise concerns on other matters.

## Impact of modular syllabuses

- 7. Please explain all the different routes through GCSE English, English language and English literature syllabuses that students could take and how well candidates on each route have performed.**

### Response

See two documents attached:

- Units available for English suite 2011/2012
- Routes taken in the English and English Language 2012

There are some units that are the same (called common units) whether candidates intend to complete an English or English language qualification, and some that are specific to the qualification being taken. All units for English literature are specific to that qualification. A list of unit titles available for each qualification is attached (Units available for English suite 2011–2012).

Centres and/or candidates are free to choose the order in which they take units, and to choose whether candidates take the higher or foundation tier of units (where available). They also have the choice to re-sit units and can choose in which series they re-sit these units. Candidates are not required to declare which of English or English language they wish to complete until their final series, and candidates may also choose to enter units specific to each of English and English language (although they will only be able to gain one qualification).

There are the following parameters: candidates may re-sit each unit only once, but they could re-take both the higher and foundation versions of units, and at least 40 per cent of the total assessment must be taken in the series in which the candidate is entered for certification (claim their GCSE), in this case June 2012.

A candidate for GCSE English or GCSE English Language, not taking any resits and not taking both higher and foundation tier options in any units, would take: three units for AQA and Edexcel English and English language qualifications and four units for CCEA, OCR and WJEC English and English language qualifications. However, a student could complete up to eight units for AQA and Edexcel, 10 units for OCR and 12 units for CCEA and WJEC (if they retook all units and entered both higher and foundation tier versions of the same written papers).

Our analysis shows that there were 2550 different combinations (routes) of units being entered and re-entered (resits) by candidates to achieve a GCSE English or GCSE English Language. There is likely to be different grade profiles for different routes because the capabilities of candidates will be different.



We have provided details of the top ten routes (aside from CCEA GCSE English where only five routes were taken), which account for 76.8% of all candidates, for each exam board, as well as the cumulative percentage of students achieving each grade for these most popular routes (Routes taken in GCSE English suite 2012).

For AQA, for example, the top 10 routes for both English and English language all involved submitting the controlled assessment units in June 2012. The differences between the routes were around when the written exam was taken, whether foundation or higher tier was entered, and whether candidates re-sat that written paper.

**8. What percentage of candidates in England who achieved grade C or above and grade A or above took one or more units before June 2012 and what percentage took all of their units in June 2012?**

**Response**

Two-thirds (67 per cent) of candidates achieving A\* C in English GCSE took at least one unit prior to June 2012. This compares with 63 per cent of the whole candidature. For the A\* A candidates, 61 per cent took at least one unit prior to June 2012, with the remaining 39 per cent taking all their units in June 2012. There were also a very small number of candidates (0.5 per cent) who took all of their units prior to June 2012 (0.5 per cent A\* A, 0.6 per cent A\* C).

We do not know that these separate groups would have achieved similar outcomes if they had all taken the same route, as these groups are self-selecting and there may be good reasons why their results vary.

	Grade		Total (Grade A-U including absence)
	A*-A	A*-C	
Candidates who sat all their units in June 2012	35886	133211	228346
% sitting all their units in June 2012	38.6	33.5	37.3
Candidates who sat at least one unit before June 2012	57064	264287	383401
% sitting at least one unit before June 2012	61.4	66.5	62.7
Total number of candidates	92950	397498	611747

#### Notes

- i) Caution should be taken when comparing the two groups (A\* A and A\* C) as the first is also part of the second group.
- ii) Based on candidate level data collected from awarding organisations (September 2012)
- iii) Covers all awards in GCSE English and English Language 2012, regardless of age and centre type

### 9. How many centres chose to put their students through GCSE English, English language and English literature early, compared with the previous year?

#### Response

It is not appropriate to compare 2012 with previous years because this is the first year of award for the new specifications and they have changed significantly.

With the change to a unitised structure, centres will, in future, be able to enter the student for the qualification in any award series (January or June), whether they are in Year 10 or Year 11. But because summer 2012 was the first full award that option was not available. Previously centres could enter students early in November or June of Year 10, if the specification offered allowed for it. However, most GCSE English specifications available in 2011 were linear. Generally, for exams, there was one series in June and one further series in November which was intended as a re-sit opportunity, and for coursework there was usually just the one series in June.

### 10. Since your initial report, what more have you learnt about the differences in entry patterns between January and June?

- a) Do some types of centre tend to enter candidates in January and/or June?

- b) Do schools/colleges enter different types of candidates in January and/or June?
- c) What impact does this have on aggregation outcomes?
- d) What have you done to try to ensure equity between exam series?

## **Response**

- a) See attached document – entry patterns for students

The tables below show the types of centres entering candidates for units prior to June 2012.

In summary the data indicates that mixed gender schools were more likely than single-sex schools to enter at least 50 per cent of candidates for units prior to June 2012.

Colleges, independent schools and special schools were more likely than other centre types to enter 50 per cent or more of their candidates for all units in June 2012. This pattern was also apparent in schools with the highest age group (19+ years).

In regions East of England, London, South East and Wales, a greater proportion of centres entered more than half of their candidates for all unit in June, as opposed to early entries. In all other regions, centres entered more than half of their candidates early.

Finally, centres with fewer than 100 candidates were more likely than other centre types to enter 50 per cent or more of their candidates for all units in June 2012.

## Candidate entry patterns for GCSE English by the proportion of candidates within a centre and centre characteristics

### Single or mixed gender centre

Gender of centre	Total		More than 50% of candidates in a centre sat at least one unit before June 2012		More than 50% of candidates in a centre sat all units in June 2012	
	Number of centres	%	Number of centres	%	Number of centres	%
Boys	239	5.4	108	4.6	131	5.6
Girls	361	8.1	138	5.9	223	9.5
Mixed	3595	80.9	1975	84.5	1620	69.3
Not applicable	249	5.6	115	4.9	134	5.7
Total	4444		2336		2108	

### Region of centre

Region of centre	Total		More than 50% of candidates in a centre sat at least one unit before June 2012		More than 50% of candidates in a centre sat all units in June 2012	
	Number of centres	%	Number of centres	%	Number of centres	%
East Midlands	322	7.2	193	8.3	129	6.1
East of England	425	9.6	188	8.0	237	11.2
London	600	13.5	289	12.4	311	14.7
North East	201	4.5	110	4.7	91	4.3
North West	593	13.3	364	15.6	229	10.8
Not Applicable	30	0.7	16	0.7	14	0.7
South East	707	15.9	308	13.2	399	18.9
South West	425	9.6	236	10.1	189	9.0
Wales (pseudo)	260	5.8	118	5.0	142	6.7
West Midlands	488	11.0	277	11.8	211	10.0
Yorkshire and the Humber	398	8.9	239	10.2	159	7.5
Total	4449		2338		2111	

### Type of centre

Type of centre	Total		More than 50% of candidates in a centre sat at least one unit before June 2012		More than 50% of candidates in a centre sat all units in June 2012	
	Number of centres	%	Number of centres	%	Number of centres	%
Academies	1296	29.1	836	35.7	460	21.8
Colleges	284	6.4	43	1.8	241	11.4
Free Schools	6	0.1	3	0.1	3	0.1
Independent schools	540	12.1	172	7.4	368	17.4
LA maintained schools	1728	38.8	1094	46.8	634	30.0
Other types	31	0.7	16	0.7	15	0.7
Special schools	305	6.8	56	2.4	249	11.8
Welsh schools	263	5.9	119	5.1	144	6.8
Total	4453		2339		2114	

### Highest statutory age of pupils at centre

Highest age of centre	Total		More than 50% of candidates in a centre sat at least one unit before June 2012		More than 50% of candidates in a centre sat all units in June 2012	
	Number of centres	%	Number of centres	%	Number of centres	%
15 years or less	7	0.2	3	0.1	4	0.2
16 Years	1435	34.2	772	34.8	663	33.6
17-18 years	2020	48.2	1202	54.1	818	41.5
19+ years	729	17.4	243	10.9	486	24.7
Total	4191		2220		1971	

## Number of candidates entered for GCSE English

Size of centre	Total		More than 50% of candidates in a centre sat at least one unit before June 2012		More than 50% of candidates in a centre sat all units in June 2012	
	Number of centres	%	Number of centres	%	Number of centres	%
<50 candidates	2281	43.7	939	35.7	1342	51.8
50-99 candidates	568	10.9	266	10.1	302	11.6
100-149 candidates	724	13.9	402	15.3	322	12.4
150-199 candidates	778	14.9	494	18.8	284	11.0
200-249 candidates	502	9.6	308	11.7	194	7.5
250+ candidates	372	7.1	223	8.5	149	5.7
Total	5225		2632		2593	

### Notes:

- i) All demographics except the number of candidates taking GCSE English in a centre have been matched in from Edubase([www.edubase.gov.uk/edubase](http://www.edubase.gov.uk/edubase)) and therefore only centres that could be matched using unique centre identifiers are included.
- ii) All centre sizes are included, including smaller sized centres.

- b) See attached document – Gender and year group by centre type

The following tables show the entry patterns for GCSE English for candidates by gender, age, free school meal status, SEN and first language.

### The proportion of candidates within a centre with free school meals

Candidate entry patterns	Free school meals groups						Total	% Total
	Low - less than or equal to 20%	% low	Medium - 20.01%-35%	% medium	High - 35.01%+	% high		
Candidates sitting all units in June 2012	136745	36.7	22990	22.2	12620	29.9	172355	33.2
Candidates sitting at least one unit before June 2012	235995	63.3	80518	77.8	29575	70.1	346088	66.8
Candidates sitting all units before June 2012	1588	.4	743	.7	298	.7	2629	.5
Candidates sitting at least one unit before June 2012 and no units in January 2012	80672	21.6	25146	24.3	8146	19.3	113964	22.0
Total candidates	372740		103508		42195		518443	

### Candidate entry patterns for GCSE English by the proportion of candidates within a centre with Special Education Needs or School Action Plus (SAP)

Candidate entry patterns	Percentage of pupils with SEN statement or on School Action Plus (SAP)				Total	% Total
	Less than or equal to 10% SEN or SAP	% Less than or equal to 10% SEN or SAP	More than 10% SEN or SAP	% More than 10% SEN or SAP		
Candidates sitting all units in June 2012	128695	35.6	44059	28.1	172754	33.3
Candidates sitting at least one unit before June 2012	232358	64.4	112989	71.9	345347	66.7
Candidates sitting all units before June 2012	1666	.5	963	.6	2629	.5
Candidates sitting at least one unit before June 2012 and no units in January 2012	80965	22.4	32654	20.8	113619	21.9
Total candidates	361053		157048		518101	

### The proportion of candidates within a centre that do not have English as a first language

Candidate entry patterns	Percentage of pupils with English not as first language				Total	% Total
	Less than or equal to 20% of candidates with English not as first language	% Less than or equal to 20% of candidates with English not as first language	More than 20% of candidates with English not as first language	% More than 20% of candidates with English not as first language		
Candidates sitting all units in June 2012	134257	32.8	35906	35.8	170163	33.4
Candidates sitting at least one unit before June 2012	275027	67.2	64393	64.2	339420	66.6
Candidates sitting all units before June 2012	2150	0.5	478	0.5	2628	0.5
Candidates sitting at least one unit before June 2012 and no units in January 2012	92329	22.6	19700	19.6	112029	22.0
Total candidates	409284		100299		509583	

#### Notes:

- Free schools meals, SEN and SAP and English not as first language data is from Department for Education School performance data. ([www.education.gov.uk/schools/performance/download\\_data.html](http://www.education.gov.uk/schools/performance/download_data.html))
- Centres are matched by local establishment metric. Only those centres that match are included.

From speaking to school leaders we understand that schools and colleges enter students in different series for a wide range of reasons that are unique to the individual circumstances of the school and the individual. These include, but are not limited to:

- entering students in earlier series to give them examination experience and enable them the opportunity to re-sit in the final series
- entering students in earlier series to enable students and teachers to concentrate teaching time on other subjects
- entering students in earlier series over concern that students will not continue in formal education up to the final series
- entering students entirely in the final series as they will be more mature at that point.

In addition evidence from analysing the routes taken by students through GCSE English and English language suggests timing relates mainly to the written paper, not controlled assessment.

Our data shows some patterns by the profile of student.

There is little variation by gender, with females very slightly more likely to have sat all of their units in June 2012. Greater variation can be seen with age. Here, candidates aged 16 and 17+ years were most likely to have completed all of their units in June. The age group which were most likely to have completed some units prior to June 2012 were those aged 15 years.

Centres with a higher percentage of students with Special Education Needs or a school action plus are slightly more likely to enter at least one unit before June 2012. There is little difference in entry pattern in centres with a higher percentage of students eligible for free school meals or with English not as a first language however.

- c) Early entry of units means that some students will have units 'banked' when they come to aggregate in the final series, in this case in summer 2012. Given the relatively small proportions of students entering early units, the results already banked did not appear to have a material effect when exam boards were aggregating in the summer.
- d) Equity, in terms of comparable demands and comparable grade standards, between series is a principle followed by exam boards and regulators. Exam boards are responsible for setting appropriate unit standards. The GCSE, GCE, Principal Learning and Project Code of Practice requires awarders to use the widest range of information available to them to ensure that the standard is maintained between series. We know from our discussions with Chairs of examiners that awarding committees took great care to align standards between series, including reviewing candidate work from the previous series

when making judgements. However, it is also clear that this was challenging in some units, particular in the 2011 units, because the standard of work was much lower than expected, generally because candidates were, in the case of January 2011, only a few months into the course.

## January and June awards

Please detail the following for candidates in England and in Wales:

- 11. For each exam board and each English syllabus (English, English language and English literature), please set out the grade boundary marks for each unit in January and June and the % of candidates achieving grades A\*-U by unit and overall.**

### Response

See tables attached GCSE and unit series grading

The tables show the percentage of candidates' achieving grades A\*-U by unit and overall, for each English syllabus (English, English language and English literature) and by board.

- 12. Overall and for each exam board, how many students took a GCSE in each of the English syllabuses in January and in June 2012?**

### Response

See attached document – Candidates sitting in January and June by syllabus in England and Wales

GCSE English and English language were new qualifications and as such were awarded for the first time in June 2012 (only unit grades were awarded before this time). In order to answer the question we have classified the January series as those candidates that have completed all of their units by January 2012. For the June series candidates, we have included all students sitting their final unit in the June 2012 window.

The data identifies the number of students that took GCSE English and GCSE English language syllabuses in January and June 2012 overall and by board.

Our analysis showed that the majority of candidates irrespective of awarding organisations and country took their final unit in June 2012 (99.6 per cent English in England, 99.5 per cent English language in England and 99.9 per cent English Language in Wales). The attached tables show that the largest majority of candidates that had completed their syllabus in January 2012 in England sat Edexcel syllabuses



with 68.9 per cent (415 out of 611) sitting English and 65.7 per cent (1529 out of 2326) in English Language.

Note: Entry patterns were not requested for English literature from the awarding organisations for manageability reasons. If the Select Committee still requires this information we will request the additional information.

**13. Overall, what number and proportion of candidates achieved grade A\*, A, B, C and D in English, English Language and English Literature in January 2012 and in June 2012?**

**Response**

See attached document – GCSE data and entry patterns for England and Wales per exam board

The data identifies what number and proportion of candidates achieved individual grades A\* U, absence grades overall and by board in GCSE English and GCSE English Language in January and June 2012.

Our analysis showed that in England and Wales (as mentioned in the previous question) the majority of candidates finished their qualification in June 2012. The results also show that in England, candidates completing in January were more likely to receive an A\* C than those in June. For English, 51.3 per cent achieved an A\* C versus 32.5 per cent in June. For English language 88.2 per cent achieved an A\* C in January compared with 76 per cent in June.

However, conclusions cannot be drawn from the differences between January and June as they are different groups of candidates.

In Wales only 39 candidates finished their qualification by January 2012 of which 34 received an absent grade and 4 were unclassified.

Note: As mentioned previously, entry patterns were not requested for English literature from the awarding organisations for manageability reasons. If the Select Committee still requires this information we will request the additional information.

**14. How do the figures for Question 13 compare with the previous year's GCSE results?**

**Response**

See attached documents – GCSE data for England and Wales 2011.

In order to compare results for 2012 with 2011 it is important to remember that the syllabuses were different across the two years for the English syllabuses.

In 2011 we saw that for England and Wales candidates, 65 per cent and 61 per cent respectively received A\* C. In 2012, for the English syllabus 24 per cent of candidates were awarded an A\* C (England only). For English language 75 per cent of candidates in England were awarded A\* C. In Wales this figure was 58 per cent.

For English literature the results were similar year on year with 76 per cent of candidates in England being awarded an A\* C grade in 2012, compared to 79 per cent in 2011. In Wales, there were 67 per cent of candidates awarded A\* C in 2012 compared to 68 per cent in 2011.

## January awards

15. **Why do you judge that the January awards were generous? What evidence is there to support this?**
16. **Why was it only possible to know retrospectively that the January awards were generous?**

### Response

We have responded to questions 15 and 16 together.

In the January 2012 units (and in 2011 units) awarders had statistical predictions for the candidates entered. These were less reliable for two reasons:

- The entry was a subset of the whole cohort and we know that the predictions are more reliable for the whole cohort than for relatively small (and possible unrepresentative) subsets
- The predictions were based on what students were expected to achieve at the end of a two year course. But in January those students had not completed a two year course, and awarders could not know how far those students had got in reaching the standard they would at the end of two years
- Awarders and exam boards 'adjusted' the statistical predictions to take account of these factors, which are difficult to quantify. It is worth noting that in making these adjustments and setting boundaries in early units, many awarders reported concerns that they were being too severe.

We have reviewed the awarding documentation from the January awards and it is clear that awarders used the predictions cautiously, expecting them to be higher than the candidates were likely to achieve. In many cases the achieved A\*-C percentage on the units was considerably lower than the predictions. It is clear from the documentation that, at the time, awarders believed they were making appropriate judgements.

Awarders had more information in June 2012 including more script evidence from a greater range of candidates, and information on matched candidates that could be made specific to the qualification they were entering.

When the summer 2012 awards were made, awarders had the benefit of more reliable statistical data for the whole cohort. With that knowledge, and having set June 2012 boundaries, it appears that some boundaries in some units in January 2012 may have been generous.

- 17. Given that it was known that the January 2012 awards would set a benchmark for schools/colleges and that some of the entrants for the January modules would be cashing in their assessments for a subject grade, what steps did Ofqual take to ensure that the exam boards made awards which met the required standard at each of the grades which were determined by judgement? What evidence were the exam boards expected/required to use?**

### **Response**

Summer 2012 was the first time these qualifications should have been cashed in. In a very small number of cases (0.5%) candidates cashed in in the summer without having sat any units (having met the requirements of the syllabus in an earlier series). This was only ever intended to be used in exceptional circumstances.

At unit level, schools and colleges who did enter students in January 2012 or earlier, would have had limited examples of student work that would have indicated what performance looked like at particular grades. Exam boards give guidance to schools and colleges about how grade boundaries are calculated, and how they can change from series to series (grade boundaries are not fixed briefing and grade boundaries over time briefing; these briefing papers are supplied for Question 25).

Exam boards follow Section 6 of the GCSE, GCE, Principal Learning and Project Code of Practice when awarding. These guidelines ensure that awarders use all the appropriate evidence to set the standard.

Further, as these were new qualifications we monitored the conduct of a sample of these unit awards in January 2012 as part of our ongoing scrutiny programme, as well as holding regular meetings with the exam boards during 2011 and 2012 to discuss the standards for new GCSE specifications, including the GCSE English suite.

- 18. What proportion of the January entry cashed in for a subject grade? On what basis did Ofqual and the exam boards conclude initially that these awards appeared to be harsh? At what stage, and on what evidence, did the view change to the awards being lenient? If this was prior to the summer grade awarding meetings, were exam boards alerted in advance to the need to tighten grade boundaries to ensure comparability with the awards of previous years?**

### **Response**

There were no subject grades awarded in January 2012. However, a small number of students completed all units required, and satisfied the terminal rule, before June 2012. We now know that there were around 3000 students, or 0.5% of the total cohort who fell into this group. Exam boards wouldn't have been able to tell before June 2012 series that these students would have completed the qualification in this way, as the exam boards had no way of anticipating whether they would re-enter units in June 2012.

So in January there was no subject aggregation, although there were unit awards. For the reasons outlined in response to Q14 and Q15, exam boards had predictions that were likely to be too high, and the awarders used their judgement, based on a review of student work, in deciding how far to adjust those predictions. When boundaries were set that produced outcomes that were much lower than predictions, awarders expressed some concerns that they had been harsh.

**19. Did Ofqual commission subject experts to observe the January awards in GCSE English? If so, please send the notes of any subject experts and/or Ofqual staff who observed the awards.**

**Response**

Yes. We observed a sample of the awarding meetings for the January 2012 exam series. This was done as part of our scrutiny programme, which is a central element of our qualifications monitoring work. The scrutiny programme is intended to help ensure that examination standards are fair, effective, reliable and consistent. It addresses a range of qualifications and exam boards each year. Recently, the scrutiny programme has tended to focus on new qualifications taken by large numbers of candidates to help check that these are being delivered effectively by exam boards.

Each individual scrutiny is an in-depth study of the assessment and awarding process for a particular qualification. Observation of exam board meetings, including awarding meetings, is just one element of the scrutiny programme.

We began scrutiny programmes in GCSE English and GCSE English literature in 2011, which will conclude in 2012. These address AQA GCSE English, Edexcel GCSE English Language and (on behalf of the CCEA regulator) CCEA GCSE English Language.

The records of January awarding meetings for AQA and Edexcel are provided. The CCEA scrutiny did not include a review of the January 2012 awarding meeting. There was, however, an observation of the OCR GCSE English Literature award from this time, despite it not being part of a formal scrutiny programme. The following 12 reports attached reflect the notes made from this process. It should be noted that

these are not normally published as they are used to inform the reports made to exam boards, and the final public report that summarises our monitoring activities.

The attachments cover reports from the observation of standardisation meetings, awarding, moderation visits and Question Paper Evaluation Committee (QPEC) meetings. We observe awarding organisation meetings, primarily not exclusively as part of the scrutiny programme, to check that they fulfil their purpose and comply with the relevant regulatory criteria. So, for example, we observe standardisation meetings to check whether examiners and moderators have obtained a clear and common understanding of how to apply the mark schemes or marking criteria. We attend awards to check that these are made based on the appropriate qualitative and quantitative evidence and, where appropriate, are in line with the relevant agreed approach to standards, for example, comparable outcomes.

- a) AQA GCSE English award (4700) February 2012
- b) AQA GCSE English examiner standardisation (ENG1F/H) January 2012
- c) AQA GCSE English examiner standardisation (ENG1F/H) January 2012
- d) AQA Moderation visits (ENG02) February 2012
- e) AQA GCSE English QPEC (ENG1F/H) January 2012
- f) Edexcel GCSE English Language award (2ENO1) February 2012
- g) Edexcel GCSE English Language standardisation of moderators (5EN03) January 2012
- h) Edexcel GCSE English Language standardisation of moderators (5EH01) January 2012
- i) Edexcel GCSE English Language standardisation of examiners (5EN2 F/H) January 2012
- j) OCR GCSE English Literature award (A661-A664) February 2012
- k) OCR GCSE English Literature standardisation of examiners (A664) January 2012
- l) OCR GCSE English Literature standardisation of examiners (A663) January 2012

As many of the documents above contain technical terms or abbreviations we have appended a glossary entitled 'roles and definitions'.

**20. Please provide the Committee with copies of your exchanges (correspondence and e-mail) with each of the exam boards between 1st December 2011 and 1st March 2012 relating to the January GCSE awards (in any or all subjects), plus notes of any meetings during that period with any or all of the exam boards about the January GCSE awards.**

**Response**

As part of our work we meet regularly with the exam boards to agree with exam board responsible officers our approach to setting and maintaining standards. We also meet with exam board technical specialists to discuss the detail of those approaches. This is the Standards and Technical Issues Group (STIG) which meets on a monthly basis.

One of the Ofqual's key activities is the data exchange of GCSE and GCE qualification outcomes every summer. This is when the regulators receive and review all provisional outcomes after exam boards' awarding processes but prior to results day. This is governed by the principles and procedures agreed between Ofqual and exam boards.

In other awarding series (such as January) there is a more qualitative "data exchange", where Awarding Organisations report on any issues they wish to draw to our attention. These reports are provided below.

The following documents are attached. AQA correspondence:

- a) Evaluative report – GCSE English Literature, January 2012
- b) Evaluative report – GCSE Maths, November 2011
- c) Evaluative report – GCSE English/EnglishLanguage, January 2012
- d) Evaluative report – GCSE Science A, January 2012
- e) Evaluative report – GCSE ICT, January 2012
- f) Evaluative report – GCSE Science B, November 2011
- g) Template for A level, AS, other GCSE report

The following documents are attached. OCR correspondence:

- h) Evaluative report for AS/A level and other GCSE Subjects
- i) Overview of all the evaluative reports.

The following documents are attached. WJEC correspondence:

- j) Evaluative report – GCSE Welsh, January 2012
- k) Evaluative report – GCSE ICT, January 2012
- l) Evaluative report – GCSE English/English Language, January 2012
- m) Evaluative report – GCSE Mathematics, January 2012
- n) Evaluative report – GCSE English Literature January 2012
- o) Evaluative report – GCSE Science 2012

The following documents are attached. Edexcel correspondence:

- p) Evaluative report for AS/A level and other GCSE subjects
- q) Evaluative report for New GCSE English, GCSE English Literature, GCSE English Language

The following documents are attached. CCEA correspondence:

- r) CCEA GCSE Awards –Spring Series 2012

The following STIG minutes are attached:

- s) STIG minutes 7th December 2011
- t) STIG minutes 18th January 2012
- u) STIG minutes 14th March 2012



## June awards

- 21. Did the January 2012 outcomes have a bearing on the June 2012 outcomes and if so, how? Were the June exams graded more harshly to compensate for the generous January awards? It has been suggested that too many grade Cs were awarded in January 2012, leaving too few available in June 2012. Is this the case?**

### Response

Unit awards are conducted using all the information available to the awarding committee, that may have included candidate work from the January series, together with statistical information on how candidates performed in previous series, as part of the evidence that awarders used to determine June unit grade boundaries, alongside the other evidence outlined in our answer to Question 17.

We concluded that some of the grade boundaries on some of the units in January 2012 were generous. Some commentators have suggested that the June boundaries were harsher than they needed to be in order to compensate for the January boundaries. The exam boards have told us that they are confident that the boundaries in June were set in the right way. We have seen the evidence from the awarding meetings that exam boards followed the GCSE, GCE, Principal Learning and Project Code of Practice in setting these boundaries.

To review results prior to subject grades being awarded exam boards and regulators use predictions that are available for a sub-set of candidates – those 16-year-olds who have been matched to their Key Stage 2 results (in this case from 2007). We call those “matched candidates”. Because the overall entry can vary from one year to the next we use the matched candidates as a “like-for-like” comparison.

We knew from the cohort’s Key Stage 2 results that they were not as strong a group as those who did GCSEs in 2010 and 2011. That the final subject outcomes for this sub-set of students (called “matched candidates”) was higher than the outcomes predicted when we applied the comparable outcomes approach is probably due to the slightly generous boundaries in some of the January 2012 units. So there is no evidence that June boundaries were harsh in order to compensate for some generosity in January. In our final report we will set out in more detail the comparisons of matched candidates with predictions, and we also explain the drop in overall results for all candidates.

There is no quota for the number of candidates that can be awarded each grade.

- 22. Were June exams graded to make them directly comparable with GCSE English exams in previous years or were they graded to make 2012 exams**

**overall comparable on average, given that the January awards were deemed to be generous?**

**Response**

The English GCSE suite of qualifications offered for the first time in 2010 were designed to be of the same standard as the previous English GCSE specifications offered until 2011. This standard is maintained at qualification level.

As the units for these qualifications were new in 2010, there is no scope to make the standard required directly comparable on a unit level. In the early series of awarding units, the evidence base, both qualitative and quantitative, is less reliable than when units have been offered for a number of years. That is why awarders use a range of evidence when making decisions on grade boundaries (see our response to Question 17). Awarders would have reflected on standards in previous series as part of their deliberations and compared them with the evidence they were seeing in the current series.

**23. Did the June 2012 written papers have to be graded more harshly to compensate for stronger performance on controlled assessment units in order to avoid grade inflation? Was this of greater impact than the difference between the January and June outcomes?**

**Response**

No. The grading process for each unit is undertaken independently. A stronger performance in controlled assessment would have led to more grade Cs for the controlled assessment units. There was no evidence of stronger performance in controlled assessment units from the evidence in reports from awarders.

Each unit was awarded separately based on the range of evidence available for each unit award in each series. Some of the grade boundaries on some of the units awarded in January 2012 were probably generous, but it was certainly not all of the units. For example, we know that the AQA Unit 1 foundation tier paper was a more challenging paper in January than it was in June, and that is why there was a different grade boundary set for each, not because of generosity in the earlier series.

That the matched candidate figures were up on predictions is likely due to the slightly generous boundaries in some of the January 2012 units. This is explored further in our response to Question 35.

**24. What were the “unknowns” in January 2012 which were known in June 2012 and why?**

**Response**

In January awarders only had Key Stage 2 predictions for those entered at each series, not their entire entry cohort. As these predictions were for a subset of the cohort they may or may not have been representative of the whole cohort. They also represented levels of achievement at the end of a two-year course but students were not at the end of their course when they entered these units. In June the Key Stage 2 predictions were for the whole cohort and most would have been at the end of their two-year course.

The awarders did not know which qualifications students would be certifying for (English or English language). This was because students could decide late in their course which qualification they would certificate for. In June students would have indicated which qualification they would enter, meaning that predictions on matched candidates could be made specific to the qualification entered.

Awarders had script evidence from a more limited range of candidates in January as this was the first award of these qualifications. In June there was script evidence from a greater range of candidates covering the whole cohort. This would provide more help to examiners when making their judgements.

**25. What evidence is there to suggest that schools/colleges shouldn't have been "taken by surprise" by the variation in grade boundaries between January and June 2012 results?**

**Response**

There are two attachments:

- Grade boundaries are not fixed briefing
- Grade boundaries over time briefing

Grade boundaries change between series. This is clear from looking at previous awards, but also through the guidance given by exam boards (grade boundaries are not fixed briefing and grade boundaries over time briefing attached).

**26. You have said that "had the grade boundaries for January carried through to June, there would have been very significant grade inflation."<sup>1</sup> How much grade inflation would have occurred overall and how was this distributed across the exam boards? How many students would have gained a grade C had the January grade boundaries been carried through to June?**

---

<sup>1</sup> GCSE English Awards 2012: A Regulatory Report (Ofqual, 2012 p16) [www2.ofqual.gov.uk/files/2012-08-31-gcse-english-awards-2012-a-regulatory-report.pdf](http://www2.ofqual.gov.uk/files/2012-08-31-gcse-english-awards-2012-a-regulatory-report.pdf)

## Response

The data below refers to all candidates if the January 2012 grade boundaries had been applied in June 2012. This demonstrates that grade inflation would have been extremely significant had all June boundaries been amended to reflect January boundaries, particularly in English specifications. This would have varied by awarding organisation, however. For example, the percentage of students gaining C or above in English would have increased by 17 per cent and 16.8 per cent for AQA and OCR English candidates respectively. Data for CCEA has not been provided as they had no candidates in England.

### WJEC

Specification	Actual % of candidates gaining a C or above June 2012	% of candidates gaining a C or above June 2012 if January 2012 grade boundaries applied
4170 (English Language)	72.7	73.8
4190 (English)	33.4	33.1

Note: WJEC did not offer controlled assessment until June 2012; therefore there were no changes to controlled assessment grade boundaries to apply in the table above.

### OCR

Specification	Actual % of candidates gaining a C or above June 2012	% of candidates gaining a C or above June 2012 if January 2012 grade boundaries applied
English Language (J355)	81.8	89.3
English (J350)	36.9	53.7

### AQA

Specification	Actual % of candidates gaining a C or above June 2012	% of candidates gaining a C or above June 2012 if January 2012 grade boundaries applied
English Language	74.8	84.2
English	31.3	48.3

**Edexcel**

Specification	Actual % of candidates gaining a C or above June 2012	% of candidates gaining a C or above June 2012 if January 2012 grade boundaries applied
English Language	75.9	79.6
English	36.7	48.8

**27. What sort of adjustments did you ask WJEC to make to its grade boundaries?****Response**

The letter from Dennis Opposs to Gareth Pierce on 8th August asked WJEC “to review the English and English language awards at grade C in order to produce outcomes that are much closer to predictions” as we had “not received any evidence to support those outcomes as appropriate”. WJEC replied with three options, and we and the Welsh Regulator agreed with WJEC’s preferred option for adjustments. For the letter exchange, please see the attachments included in the response to Question 29.

**28. Did Ofqual at any point ask (or consider asking) AQA or OCR to reconsider their grade boundaries?****Response**

We did not at any point ask or consider asking AQA or OCR to reconsider their grade boundaries. Their grade outcomes at subject level were within tolerance and we saw no reason to challenge their outcomes.

**29. Please provide the Committee with all exchanges and notes of meetings between you and the exam boards about the summer GCSE awards between 1 June 2012 and 1 September 2012.****Response**

As also referred to in question 20, as part of our work we meet regularly with the exam boards to agree with exam board responsible officers our approach to setting and maintaining standards. We also meet with exam board technical specialists to discuss the detail of those approaches. This is the Standards and Technical Issues Group (STIG) which meets on a monthly basis.

One of the Ofqual's key activities is the data exchange of GCSE and GCE qualification outcomes every summer. This is when the regulators receive and review all provisional outcomes after exam boards' awarding processes but prior to results day. This is governed by the principles and procedures agreed between Ofqual and exam boards.

In other awarding series (such as January) there is a more qualitative "data exchange", where Awarding Organisations report on any issues they wish to draw to our attention. These reports are provided below.

Attached is WJEC correspondence:

- a) WJEC Summer 2012 GCSE English award (8 Aug 2012)
- b) WJEC – Summer 2012 GCSE English and English Language boundary marks (9 Aug 2012)
- c) WJEC – Summer 2012 GCSE English and English Language (9 Aug 2012)

Attached is Edexcel correspondence:

- d) Ofqual letter to Edexcel – Summer 2012 GCSE English award (7th August 2012)
- e) Ofqual letter to Edexcel – Summer 2012 GCSE English award (9th August 2012)
- f) Ofqual letter to Edexcel – Summer 2012 GCSE English award (10th August 2012)
- g) Edexcel letter to Ofqual – Summer 2012 GCSE English award (8th August 2012)
- h) Edexcel letter to Ofqual – Summer 2012 GCSE English award (10 August 2012)
- i) Edexcel to Ofqual email trail – Summer 2012 GCSE English award (9 August 2012)

Attached is correspondence relating to GCSE Science:

- j) Ofqual letter to AQA – Setting and maintaining GCSE standards meeting 2012
- k) Ofqual letter to AQA – GCSE Science Summer 2012

- l) Ofqual letter to Edexcel – Setting and maintaining GCSE standards meeting 2012
- m) Ofqual letter to Edexcel – GCSE Science Summer 2012
- n) Ofqual letter to OCR – Setting and maintaining GCSE standards meeting 2012
- o) Ofqual letter to OCR – GCSE Science Summer 2012
- p) Ofqual letter to WJEC – Setting and maintaining GCSE standards meeting 2012
- q) Ofqual letter to WJEC – GCSE Science Summer 2012
- r) Response AQA to Ofqual
- s) Response AQA to Ofqual
- t) Response Edexcel to Ofqual
- u) Response OCR to Ofqual

Attached is correspondence relating to Maintenance of standards meetings:

- v) Letter to AQA – Maintenance of GCSE and GCE standards Summer 2012
- w) Letter to CCEA – Maintenance of GCSE and GCE standards Summer 2012
- x) Letter to Edexcel – Maintenance of GCSE and GCE standards Summer 2012
- y) Letter to OCR – Maintenance of GCSE and GCE standards Summer 2012
- z) Letter to WJEC – Maintenance of GCSE and GCE standards Summer 2012

Attached is CCEA correspondence:

- aa) CCEA letter to Ofqual – examination outcomes, July 2012
- bb) Ofqual letter to CCEA – Examination outcomes, June 2012

Attached are STIG minutes and documentation tabled at the meeting on June 2012:

- cc) Item 2 – Minutes May 2012 STIG meeting

- dd) Item 3 – AQA Summary of June 2012 predictions
- ee) Item 3 – Edexcel Summary of June 2012 predictions
- ff) Item 3 – GCSE Science 2011 15 year olds (doc. 1) as at August 2011
- gg) Item 3 – GCSE Science 2011 15 year olds (doc. 2) as at June 2012
- hh) Item 3 – OCR Summary of June 2012 predictions
- ii) Item 3c – Procedures for summer 2012 data exchange
- jj) Item 5 – STIG draft terms of reference
- kk) Technical seminar proposed agenda
- ll) June 2012 Agenda
- mm) June 2012 minutes of meeting

Attached are Data Exchange teleconference minutes from summer 2012:

- nn) Teleconference notes 06.07.12
- oo) Teleconference notes 13.07.12
- pp) Teleconference notes 17.07.12
- qq) Teleconference notes 20.07.12
- rr) Teleconference notes 27.07.12



**30. Please supply the Committee with copies of the reports from staff and subject experts who observed GCSE English awarding meetings in any of the exam boards for the summer 2012 awards.**

**Response**

Attached are the following three reports from summer 2012:

- a) AQA – GCSE English award (4700) July 2012
- b) CCEA – GCSE English /English Language (G9290) July 2012
- c) Edexcel – GCSE English Language award (2EN01) July 2012

(Note: the Welsh Government carried out scrutiny for WJEC in 2012, and there was no scrutiny for OCR English awards in 2012)

**31. Can you confirm that the statistical outcomes of the summer 2012 English award related solely to the candidates who were cashing in their assessments for a subject grade in June 2012 – i.e. the statistical picture did not include those candidates who received their subject grade in January?**

**Response**

No candidates received their subject grade in January – the first series for certification for these qualifications was June 2012. However there were a small number of candidates that completed their units before June 2012 but only received their award in June. Those candidates will have been included in the overall results issued in August 2012 and will have been included in any data reviewed as part of the award of subject grades in the summer award meeting.

**32. How many students in England were affected by the changes made to Edexcel's grade boundaries as a result of Ofqual's intervention at each grade A\* U?**

**Response**

We challenged Edexcel to justify its preliminary results for GCSE English, as they were outside the tolerance agreed between the regulators and the exam boards. We did not intervene to instruct Edexcel to move any particular grade boundaries. Our exchanges of letters (as attached as part of our response to Question 29) occurred alongside internal processes that were on going within Edexcel, who were already considering similar issues. At the conclusion of the exchange Edexcel informed us that they were intending to move the grade boundary on Unit 5EH03 from 62 to 65, which meant that while the end results were still outside of tolerance they were much more in line with the predictions. We have asked Edexcel to model these changes to

answer this question, but to date (17th October) they have not provided us with this information.

**33. What alternative courses of action have you considered with regard to the summer 2012 GCSE English results and why did you reject them?**

**Response**

We will be making recommendations to our Board ahead of publication of our final report by the end of October, and our report will provide further details of the options we have considered and the rationale for the Board's decision.

**34. What advice have you received from your Standards Advisory Group regarding the 2012 GCSE English results and the action you have taken? What advice has the Group given you in connection with next year's awards?**

**Response**

The Standards Advisory Group considered the matter of grade inflation in March 2012 and established a sub group to work on options for the approach to this in the future. In September 2012 the group met shortly after Ofqual had published its interim report and had a lengthy discussion on summer 2012 GCSE English grading and Ofqual's interim report. They explored the factors that may have combined to cause the outcomes and went on to explore any changes that could be made to achieve robust awarding in 2013.

## Comparable outcomes

### 35. What modelling did you do to test the likely effect of the comparable outcomes approach on results in GCSE English?

#### Response

“Comparable outcomes” is an approach to setting grade standards for the whole cohort. The prediction matrices are developed by the exam boards based on a sub-set of candidates – those 16-year-olds that have been matched to their Key Stage 2 results (in this case from 2007). We call those “matched candidates”. Because the overall entry can vary from one year to the next we use the matched candidates as a “like-for-like” comparison. This approach, agreed between the regulators and exam boards, has been to use a combination of statistical predictions and senior examiner judgement to maintain standards between old and new qualifications. The purpose of the comparable outcomes model is to provide statistical input to the awarding meetings and to identify if any proposed award would be materially out of tolerance with predictions. It does not impose an outcome. Decisions are taken by the exam boards but they have to justify to us if they propose to award out of tolerance.

There are challenges in applying the comparable outcomes approach in early units to a sub-set of the cohort, regardless of the size of that cohort. And there were particular challenges for GCSE English and English language as it was not known until the final series which particular qualification of the two candidates would be certifying in June 2012. In developing the prediction matrix for GCSE English and English language, it was originally anticipated that this would be a common prediction for both subjects. However, when reviewing this it became clear that a single figure for each exam board was not fit for purpose because of the different ability of those taking English and English language. This is minuted in the “Summer 2012 data exchange – Teleconference between regulators and awarding organisations of 13th July 2012” attached as part of the response to Question 29. Two separate predictions were then produced. These figures, alongside the outcomes for matched candidates are detailed the table below:

	AQA % A*ꠁC	Edexcel % A*ꠁC	OCR % A*ꠁC	WJEC % A*ꠁC	Overall % A*ꠁC
<b>English language</b>					
Predicted	77.5	77.5	79.0	77.3	77.6
Actual	77.5	78.7	80.1	77.7	77.8
Difference	0	+1.2	+1.1	+0.4	+0.3
<b>English</b>					
Predicted	31.3	34.9	33.6	36.3	33.3
Actual	32.5	41.5	34.9	38.3	35.3
Difference	+1.2	+6.6	+1.3	+2.0	+2.1
<b>Combined</b>					
Predicted	67.3	65.0	67.8	65.3	66.7
Actual	67.6	67.8	68.9	66.2	67.4
Difference	+0.3	+2.8	+1.1	+0.9	+0.7

Source: data supplied to Ofqual by AQA on behalf of AQA, Edexcel, OCR and WJEC,  
31st July 2012

**36. How do you distinguish grade inflation from genuine underlying improvement? Under what circumstances would rising pass rates be legitimate, using the comparable outcomes methodology?**

**Response**

See attached document – summary of ‘Out of tolerance’

That all other things are equal is one of the principles of the comparable outcomes approach. If improved performance has genuinely been seen, then we would expect that results would increase, and that exam boards would be able to provide sufficient evidence to justify an award that was out of tolerance with predictions.

Genuine improved performance would be signalled by the awarding committee when they recommended a grade boundary. Awarding committees (and moderators as applicable) review samples of student work as part of the awarding process. If the samples indicated a higher (or lower) level of performance was being seen than would be suggested by other sources of evidence (such as statistical indicators) then awarders have the responsibility to ensure that this is reflected in the grade boundaries given. When such a change is outside of tolerance agreed with us, exam boards report these differences to us to justify this variation.

For example, of the 241 GCSE subject awards made in summer 2012, 41 were reported to us to have differed from expected outcomes and were appropriately justified by evidence. For more information see “Summary of ‘*Out of Tolerance*’ (OTT) GCSE reporting 2012”.

**37. How do you guard against a centre's grades going down because other schools/colleges have improved, even though within the centre the performance is similar to previous years (i.e. they suffer a relative decline even though they are doing as well as before in absolute terms)?**

**Response**

Variation of results within a centre between years is normal as all groups of students are different, and the overall student profile of centres can change between one year and another.

Awarding organisations award at a 'total entry' level. It would be inappropriate to adjust awarding to take account of individual schools' and colleges' results as exam boards would not know the possible reasons for any such variations.

We do not monitor on a centre by centre level, but at cohort level. Monitoring at centre level would be impractical and problematic for a number of reasons. For example: statistical indicators become less reliable when used on small groups of students, so it would not be appropriate to track at a centre level.

**38. Given the changes in entry patterns and the difficulty in predicting overall changes in outcomes, how reliable was the comparable outcomes approach as a basis for awarding decisions in GCSE English in June 2012?**

**Response**

We have discussed this approach with experts in the field and there is general agreement that, while not perfect, the comparable outcomes approach offers a means of maintaining comparable standards when qualifications change. In new qualifications, there is often a dip in performance in the first year, and the comparable outcomes approach aims to ensure that, so far as possible, these students are not advantaged or disadvantaged because they are part of the first cohort taking a new qualification.

Given the nature of the changes to English – two different qualifications, modular instead of linear, controlled assessment instead of coursework, and changes to the content and skills being assessed – setting appropriate standards proved more challenging than anticipated. However, we are confident that without it there we would have seen significant grade inflation which was not justified by improved performance, and inconsistent standards between exam boards.

**39. At what level does the comparable outcomes approach work? Is it based on national Key Stage 2 data or does it drill down to the performance of individual schools/colleges or pupils, as some have suggested? How reliable is the Key Stage 2 data?**

**Response**

See attached document – predicting GCSE outcomes

It is based on national Key Stage 2 information available for the cohort of candidates taking a particular subject. It is not used at individual candidates or schools/colleges. Such indicators work most effectively where there are very large numbers of candidates.

Predictor matrixes based on Key Stage 2 data have been modelled on a wide range of GCSE specifications and have been found to be accurate. Further details of the work carried out can be seen in the attached report “Predicting GCSE outcomes based on candidates’ prior achieved Key Stage 2 results”. But could be said that (as per Question 36) of the approximately 241 subject awards made for GCSEs in summer, 41 did not meet the predictions at award, and these could be justified. This indicates that overall, comparable outcomes were in line with actual subject outcomes, except where there were good reasons for them not to be.

**40. What additional statistical evidence was available to help define outcomes e.g. sub-cohort data such as relative performance by school type or year group or with other subjects?**

**Response**

Exam boards develop their own range of statistical tools that are appropriate for their groups of centres that are used as part of the awarding process. Ofqual does not mandate any particular tools. All candidates are treated equally at the point of awarding, with no reference to contextual factors such as type of school or age of candidate.

**41. Is the comparable outcomes approach at risk of being discredited by what has happened with GCSE English? What alternatives could be used instead?**

**Response**

As discussed in the answer to Question 38, the comparable outcomes approach operates at overall qualification level, offers a means of ensuring that students in one year receive the same result as they would have in previous years, all other things being equal. It has been one of the tools we have used since 2009 to maintain

standards in new qualifications – new AS and A levels, and in summer 2011 new GCSEs.

We have consulted with experts and still believe this is the best approach we have to maintaining standards and preventing grade inflation, although we recognise that it is not perfect. Alternatives to its use would be to allow the proportion of students getting A and C to steadily increase in spite of no evidence of increased performance (grade inflation) or to limit the proportion of students able to achieve certain grades irrespective of performance (closer to a norm referencing approach).

**42. How much use of candidates' work, as well as grade descriptions and archive material, was made to confirm that the grade boundaries selected gave comparable outcomes in terms of what students could demonstrate, as compared to statistical indicators? Are you satisfied that, at the grade boundaries selected in June 2012 by each of the exam boards, the work demonstrated by students was of the right standard? How robust do you think that evidence is?**

## **Response**

We observed a sample of the awarding meetings for the January 2012 exam series. This was done as part of our scrutiny programme, which is a central element of our qualifications monitoring work. The scrutiny programme is intended to help ensure that examination standards are fair, effective, reliable and consistent. Each individual scrutiny is an in-depth study of the assessment and awarding process for a particular qualification. Observation of exam board meetings, including awarding meetings is just one element of the scrutiny programme.

We began scrutiny programmes in GCSE English and GCSE English literature in 2011, which will conclude in 2012. These address AQA GCSE English, Edexcel GCSE English Language and (on behalf of the CCEA regulator) CCEA GCSE English Language.

From the evidence we have seen from the notes from those awarding meetings, reports from our own observers at some of these meetings, and from formal meetings we have now held with key members of the awarding committees, we are satisfied that for each awarding meeting an appropriate balance was used. Notes from the meetings demonstrate that where awarders wanted more information (such as additional scripts) they were given this information for consideration.

We are satisfied that they used their best professional judgement to arrive at the boundaries. We carry out review programmes following awarding, such as scrutiny and comparability work, to monitor that this has been done effectively.

We also have emerging evidence from our review of candidate work as part of the scrutiny that the work being produced by candidates at the A and C boundaries was of an appropriate standard.

**43. Are you satisfied that the work of candidates just below the grade C boundary could be demonstrated not to meet the requirements for a grade C when compared to any exemplar material or statement of standards?**

### **Response**

The awarding meeting specifically looks at candidates' work above and below the proposed grade boundaries to ensure the boundary is set correctly.

Yes. We are satisfied that exam boards correctly applied the Code of Practice at unit level, and, at qualification level these aggregate to the performance described by grade descriptors.

However, it should be noted that for every student that has a subject level award, there are a series of complex judgemental decisions based on the unique performance of a candidate at question, unit and qualification level that would need to be considered against grade descriptors at subject level.

And when comparing to exemplar work, it is highly unlikely that two candidates will demonstrate the same strengths and weaknesses in a particular assessment making this, also, a qualitative indicator, rather than there being a yes or no answer.



## **Controlled assessment and moderation**

- 44. What steps did awarding bodies take to (a) ensure that teachers were well prepared to assess candidates' work, (b) moderate the assessments and (c) what feedback did schools/colleges get?**

### **Response**

See attached briefing – how internal marking works

GCSEs are marked through the application of a mark scheme to a question paper or assessment by trained markers or examiners. The attached paper explains this in greater detail.

- 45. What steps were taken to ensure that teachers understood that the grade boundaries set in January would not automatically be carried forward to the June examination?**

### **Response**

Information about grade boundary setting was disseminated to centres in the guidance given by exam boards.

Details of this guidance are provided in the attachments for Question 25 (grade boundaries are not fixed briefing). This provides examples of guidance given from AQA, OCR, Edexcel and WJEC to teachers explaining that grade boundaries are not fixed, and that the raw mark required to gain a particular grade will vary depending on the examination series.

- 46. Some schools have said that they were advised by exam boards that the controlled assessment grade boundaries were unlikely to change. What guidance will you be asking exam boards to issue to ensure that this does not happen again?**

### **Response**

We have not received any evidence that this was the case. Exam boards do not accept that schools were advised that grade boundaries were unlikely to change. However, there clearly has been some confusion here and we will require each exam board to communicate its approach effectively to its centres in future.

- 47. What reassurances are you giving teachers that changes to controlled assessment grade boundaries will be properly communicated in future?**

### **Response**

As Question 46.

**48. Was there any change in the relationship between candidates' ability and their controlled assessment scores between January and June 2012?**

**Response**

There were changes in patterns of controlled assessment marks between the January and June series. The mark distributions for January and June awards for English show marked difference around the grade C boundary. There was bunching around the grade C boundary in the June graphs. There was also similar bunching around the grade boundaries seen in the June distributions for English language although the difference was not so marked as for English.

**49. Your initial report states that for AQA “the majority of folders for controlled assessment were within the marking tolerance but had been over-marked by teachers, particularly at the grade C/D borderline.”<sup>2</sup> How widespread was this “significant teacher over-marking”?**

**Response**

We are collating information on this issue from various sources, including from schools and colleges, and that this will be included in our final report. This was a comment made in one of the awarding documents supplied by AQA based on evidence seen by their moderators. This will have been documented by their moderation teams.

**50. What evidence do the exam boards and Ofqual have that supports the allegation that schools/colleges have been getting increasingly over-generous (or maximising their results) in their assessment of students' work? Is there analysis from previous English syllabuses that shows an increasing disparity between written papers and teacher assessments?**

**Response**

Analysis carried out by awarding bodies suggests that previous knowledge of grade boundaries may have influenced marking. Our interviews with teachers have found that schools target their resources (teaching time, intervention strategies) to maximise their students' chances of achieving a grade C.

Our analysis also shows that centres that had marks scaled after moderation had marked generously and those within the marking tolerance were more likely to be

---

<sup>2</sup>GCSE English Awards 2012: A Regulatory Report (Ofqual, 2012 p13) [www2.ofqual.gov.uk/files/2012-08-31-gcse-english-awards-2012-a-regulatory-report.pdf](http://www2.ofqual.gov.uk/files/2012-08-31-gcse-english-awards-2012-a-regulatory-report.pdf)

generous than harsh. We are still reviewing the moderation data (some of which is paper based rather than electronic) and will report in more detail in our final report.

QCAs report on coursework 2005 shows little evidence of marking to the upper limit of tolerance in coursework. However in most components there was a tendency for the mean mark from the centre to be slightly higher than that of the moderator. They also analysed mark distribution in relation to grade boundaries across a range of qualifications and little evidence was found of marking to grade boundaries except in GCSE English and art and design.

Ofqual's Reliability Compendium (chapter 9 '*A focus on teacher assessment reliability in GCSE and GCE*' page 393) suggests that there is little evidence of internal standardisation taking place in centres. Internal standardisation is the process by which assessors within schools (marking GCSE controlled assessment for example) ensure that their judgments are fair, valid and consistent with other internal assessors. Processes include internal standardisation meetings and monitoring of marking by a lead assessor.

**51. Why did the exam boards not deal with teacher over-marking through better moderation rather than by changing grade boundaries?**

**Response**

Moderation works by setting a tolerance around the mark that is acceptable to the exam board. The tolerance that has been used for controlled assessment moderation for several years is 6%. We are reviewing whether that was sufficient in the new GCSE English and English Language units.

**52. What is the marking tolerance which is allowed before a moderator adjusts marks? Why is there a tolerance?**

**Response**

The tolerance is 6 per cent of the marks available. This figure is agreed by the awarding organisations. The tolerance level is designed to reflect the reality that two markers will not agree exactly when judgements are qualitative, and allows for the prioritisation of those centres where the quality of marking is most in question.

When we have completed our review of the data from the summer moderation, we will consider whether controls should be tightened for future series.

**53. How much scaling and moderation was applied to the June 2012 controlled assessments? How does this compare to the scaling and moderation in 2011 and 2010?**

**Response**

In discussions that we have had with Chairs of Examiners, there were no issues around the volume of out of tolerance centres with regard to controlled assessment for these new specifications in June 2012 when compared with (the limited) previous series for these new specifications or when compared with the previous specifications. However, direct comparisons with previous years may not be reliable due to the change in the nature and proportion of assessment that controlled assessment represents when compared with coursework in previous specifications.

We have requested details from exam boards, and are awaiting detailed responses. We can supply these to the Select Committee when they are available. To date (17 October) they have reported that the percentages of centres that had their marks adjusted in June 2012 units were as follows:

**Edexcel**

5EH03 Creative responses 18 per cent

5EN03 The spoken language 17 per cent

5EH01 English today 12 per cent

**OCR**

A641 Reading literary texts 7 per cent

A642 Imaginative writing 3 per cent

A652B Spoken language 2 per cent

A651 Extended literary text and imaginative writing 1 per cent

**WJEC**

4193 English in the world of the imagination and 4194 Speaking and listening 5 per cent

4173 Literary reading and creative writing and 4174 Spoken language 8 per cent

**54. While exam paper grade boundaries may need to move slightly from one series to the next, how is it possible that controlled assessments using**

**the same titles and the same mark schemes would also need to be changed?**

**Response**

Grade boundaries vary between exam series – both within a year (for example, between January and June) and also across different years. In the case of controlled assessment where the task is common between series within the year, these changes arise from differences in marking and moderation between series, statistical information relating to the students taking the unit in the series (not available at moderation) and the professional judgement of the awarding committee based on the evidence of student work alongside archive scripts.

**55. What are the moderation arrangements for speaking and listening controlled assessment?**

**Response**

There are no specific requirements for the moderation of speaking and listening set out in the GCSE, GCE, Principal Learning and Project Code of Practice, and procedures vary between exam boards. Centres will normally receive visits from a moderator where advice is provided on the assessment of speaking and listening. The moderator may also ask to see a range of student performances as part of their visit. A third of centres are visited for moderation every year, therefore centres are usually visited once every three years.

Centres must keep the Candidate Record Forms of every student entered for the examination under secure conditions and be prepared to show this to the visiting moderator or post to a moderator if required.

Additional visits may be triggered if there is any doubt about a centre's accuracy in oral assessment (for example if results are out of line with student performance in other units).

We have met the Chairs of Examiners as part of our investigation into GCSE English and English language results and they have told us that these arrangements are not always effective in securing the standard for the speaking and listening components.

We are reviewing the effectiveness of the moderation arrangements and we will be considering whether any of the controls need to be tightened for future series.

## Risk management

**56. What risks were reported to the Ofqual Audit Committee and Board between August 2011 and September 2012 relating to the January and summer 2012 GCSE awards? Did English syllabuses specifically feature in the risk register and those reports?**

### Response

Our Board met eight times between August 2011 and September 2012 and reviewed and considered risks at each meeting. The Audit and Risk Committee met six times between these dates and reviewed risks at each meeting to provide the Board with an assurance that our risk management processes were operating to an appropriate standard.

Our corporate risk register primarily describes generic rather than specific risks to assist the Board determine our strategy and focus of attention. This review and consideration is supported by specific reports from our Standards and Research and Regulation Directorates on current, proposed and anticipated activity.

Although the corporate risk registers considered by the Board did not specifically feature the 2012 English syllabuses, they did include a number of risks covering various dimensions relating to the 2012 examinations. These included inconsistent standards, significant system failures, the use of qualifications for accountability purposes and the possibility of difficulties arising in awarding organisations.

In addition the October 2011 and December 2011 Standards and Research reports included English in wider reports on controlled assessment and GCSE reform respectively, and the May 2012 and July 2012 Standards and Research and Regulation reports included English in respect of the 2012 examination series. The September 2012 Board included a specific agenda item on GCSE English.

The Standards Directorate risk register included a risk that “the perception that Ofqual is not fulfilling its remit in relation to the maintenance of grading standards for summer 2012 GCSEs and A levels negatively impacts on public confidence”. There was not an expectation that GCSE English would be any more of an issue than other new GCSEs being awarded in summer 2012.

**57. What are the relative risks of using January’s standards or June’s standards? Did you consider whether the risk of impacting students’ life chances outweighed the risk of devaluing the GCSE or allowing students into HE or employment without the requisite skills?**

We and our Board have considered carefully the issues of fairness and, by implication, the risks of re-grading using January (to re-grade June) or June (to re-

grade January) grade boundaries for controlled assessment units. We have a statutory objective to maintain standards in qualifications. As GCSE qualifications are modular and two years in duration, we need to look beyond the relative performance of students in January and June 2012. So for example, some Year 10 pupils (those in their first year of GCSE study) sat modules in June 2012, and it is important that their results from this summer can be compared consistently with their future modules and with the performance of other students taking the modules later. We also know that there is some cross-over with students sitting units in January and re-sitting in June 2012.

There are four possible courses of action that we have considered: change grade boundaries for June to match January (the most common suggestion); change January grade boundaries to match June (this has not been suggested to us as the right course of action and we have earlier considered and rejected this option in view of the possible detriment to those students); change June grade boundaries to a lesser extent, or make changes to both June and January boundaries.

Our emerging findings suggest that changing grade boundaries in any way will not address fairness, because we are not able to identify the specific students who can be said to have been treated unfairly. Significant factors relating to the qualification design and to behaviours associated with the pressures of accountability do not affect outcomes for an identifiable group of students, but can have affected the outcomes for any student.

It has been suggested that to be fair, we should direct that June grade boundaries should be changed, to match January grade boundaries. In our final report we will set out the effect on outcomes nationally, should we direct June grade boundaries to match January grade boundaries.

**58. With so many changes happening at once, what safeguards are in place to disentangle all the elements and keep a clear national standard?**

**Response**

Our comparable outcomes approach aims to minimise any advantage or disadvantage to students of being the first to sit new qualifications. We know that maintaining standards is most difficult when qualifications change, and there has been a great deal of change in the English suite – a move from coursework to controlled assessment, and a higher weighting (60 per cent), a move from linear qualifications to modular qualifications where students can re-sit units during their two-year course, and significant changes to the content, with a choice between English (covering both language and literature) and English language. Our approach, agreed between the regulators and exam boards, has been to use a combination of

statistical predictions and senior examiner judgement to maintain standards between the old and new qualifications.



## Ofqual investigation

- 59. If your further investigations reveal that students have been deprived by the system of grades their work deserved, will you be prepared to reverse your decision not to re-visit grade boundaries?**

### **Response**

In considering the issue of fairness to all students (as opposed to fairness to particular groups of students) we have discussed the option to re-grade some or all units. Having considered all the evidence, it is our view that it would be impossible to identify those students that have been genuinely disadvantaged by the setting of particular grade boundaries. There are many reasons why students do not get the results they might have wanted, and we cannot disentangle those for GCSE English students this summer.

- 60. Why would a re-grading of boundaries of GCSE English and English Language not be considered to be fair?**

### **Response**

Please refer to our responses to Questions 57 and 58.

## Possible independent inquiry

**61. Has the Ofqual Board considered whether to commission an independent inquiry into why the problems with GCSE English occurred this year and which organisations (if any) were at fault, similar to the inquiry commissioned into National Curriculum Test delivery failure in 2008? If they have considered this, what were their reasons for deciding not to commission such an inquiry?**

### Response

Our Board has not considered commissioning an independent inquiry.

We have been taking the actions that we think necessary as a regulator, to get to the root of the concerns expressed about GCSE English results this year. We have been giving this work absolute priority and expect to report by the end of October. We believe this to be the right action for the regulator to take. In particular, we took the view that immediate action was essential.

Our Board is aware of concerns about partiality. We have responded promptly to all freedom of information requests, demonstrating our openness and commitment to transparency. We have met with representative bodies, to hear in detail their concerns about GCSE English results. We have employed independent consultants with education experience to design our case study work with schools and colleges, to conduct the interviews with individual members of staff in schools and colleges, and to report on the findings from those case studies.

Our final report will detail the ways in which we have regulated exam boards (alongside the other regulators). We will report on any shortcomings we can see in our actions, as part and parcel of the final report.

Our Board has taken the view that these are the immediate priorities.

## Relationships with Wales

- 62. WJEC is faced with contradictory instructions from the Welsh Government and the English regulator. On what grounds could Ofqual fairly consider excluding WJEC from the market of English schools/colleges if they follow the instructions from Welsh Ministers?**

### **Response**

We can require WJEC to make sure the standards set in GCSEs wherever they are taken are consistent with those of other exam boards. If WJEC fail to do this for some or all students, for example those in Wales, we can take regulatory action including withdrawal of recognition which would effectively remove them from the GCSE market in England.

- 63. How is it fair that a Welsh student living on one side of the Welsh/English border can be given one grade, while a student in England with the same score in the same exam is given a lower grade?**

### **Response**

It is not fair.